Chapter 10

Topology of Viral DNA

Javier Arsuaga*, Joaquim Roca† and De Witt Sumners‡

1. Introduction: The Organization of Double-Stranded DNA in Bacteriophage Capsids

Chromosome organization plays a key role in many biological processes. In viruses genome organization is essential for the packaging and releasing of the genome as well as for maintaining the stability of the viral capsid. This organization varies across different families and is highly dependent on the virus morphogenetic pathway (Casjens, 1997). In this review we will focus on the packing of double-stranded DNA (dsDNA) in bacteriophages. Understanding how dsDNA is packed in bacteriophage capsids is important because it arguably yields the simplest example of a full genome organization in a biological system. Furthermore bacteriophages are believed to pack their DNA similarly to some animal viruses such as herpes- and adeno-viruses (Casjens, 1997). Therefore understanding the basic principles of DNA packing in bacteriophages may provide essential insights into the viral assembly pathways for a wide range of bacterial and animal viruses.

During bacteriophage morphogenesis a proteinic procapsid is first assembled (reviewed in Johnson and Chiu, 2007; Earnshaw and Casjens, 1980; Dockland, 1992; and in Wang *et al.*, 2006). This assembly is followed by the active transfer of one copy of the viral genome from the host cytoplasm to the procapsid through a molecular motor called the

*Department of Mathematics, San Francisco State University, San Francisco, CA 94132.
†Instituto de Biologia Molecular de Barcelona, CSIC, Barcelona, Spain.
‡Department of Mathematics, Florida State University, Tallahassee, FL.

connector (Fuller *et al.*, 2007; Ibarra *et al.*, 2000; Rishov *et al.*, 1998). This packaging reaction is performed against strong repulsive forces due to the excessive bending and the self-repulsive nature of the double helix (Tzlil *et al.*, 2003; Kindt *et al.*, 2001; Smith *et al.*, 2001). Inside the capsid, the viral DNA reaches near-crystalline density (Earnshaw and Casjens, 1980) with concentrations of DNA fibers of up to 800 mg/ml (Kellenberger *et al.*, 1986) and an osmotic pressure of 50 atmospheres (Evilevitch *et al.*, 2003). Once the DNA is packed inside the capsid, the phage tail is assembled at the connector site and a mature virus is completed. How the DNA is folded under the extreme physical conditions described above is unknown. Biochemical and structural analyses have shown that DNA inside the capsid is kept in its B-form (Aubrey *et al.*, 1992), that the DNA forms local domains of parallel fibers that are 25 Å away from each other (Earnshaw and Harrison, 1977; Earnshaw *et al.*, 1978; Lepault *et al.*, 1986). Moreover, they show that there is no correlation between DNA sequences and their spatial location inside the capsid, with the exception of the DNA ends in some viruses (Chattoraj and Inman, 1974). The question whether there are contacts between the double helix and the interior of the capsid remains open. Work by Serwer and colleagues showed that such contacts were inexistent or minimal (Serwer *et al.*, 1992). On the other hand, new asymmetrical reconstruction techniques of frozen samples have revealed frequent contact points. Whether these contact points have any role in the packing or if they affect the overall geometry of the chromosome remains to be explored (Jiang *et al.*, 2006). Most studies agree on the existence of concentric layers of DNA fibers organized in spooling, toroidal or parallel fashion. Next we describe the proposed models for the long-range organization of the DNA inside the viral capsid.

1.1. *Spooling Models*

There are two basic spooling models: coaxial and concentric. The coaxially spooled model (Fig. 1) was initially proposed by Richards *et al.*, 1973 after performing electron micrographs in bacteriophages P2, λ, T4, T7 and T5. In this model, the DNA is organized in shells that are wound about an axis. The orientation of the axis has been a matter of debate and may be dependent on the virus. It has been proposed that the spooling axis can be perpendicular (Cerritelli *et al.*, 1997; Geller and Davis, 1964), parallel (Earnshaw and Harrison, 1977; Jiang *et al.*, 2006), or tilted a few degrees

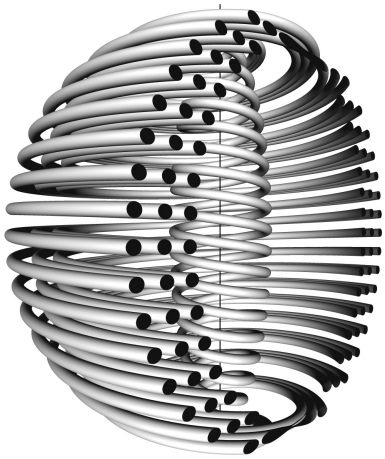

Fig. 1. Coaxial spooling model. Recent experimental results in bacteriophage phi 29 have estimated that there are about five concentric shells of DNA fibers containing up to 50% of the entire genome. This arrangement is more disordered in the center of the capsid where the radius is much shorter than the persistence length of the DNA molecule (Comolli *et al.*, 2007). The idealized model in the figure shows four coaxial shells.

(Sewer, 1986) with respect to the axis defined by the tail. In concentric models, each DNA shell is organized around an axis whose orientation is different from the previous shells. This configuration has been observed in molecular dynamics studies and represents energy minima in the absence of long-range interactions (LaMarque *et al.*, 2004).

1.2. *Toroidal Models*

Toroidal models have been proposed not only for bacteriophages but also for DNA molecules exposed to condensing agents (reviewed in Hud and Vilfan, 2005). In these models DNA travels along the surface of a torus (Fig. 2). Several toroidal models have been proposed. These include the interwound toroid (Earnshaw *et al.*, 1978), toroidal winding (Kosturko *et al.*, 1979), folded toroid (Hud, 1995) and twisted toroid (Petrov *et al.*, 2007). In the folded toroid model the DNA first assumes a conformation similar to the one illustrated in Fig. 2. Since the DNA molecule is too large to fit in the capsid by this arrangement alone the toroid needs to fold in half (or quarters) in a next step. The twisted toroid has been recently proposed

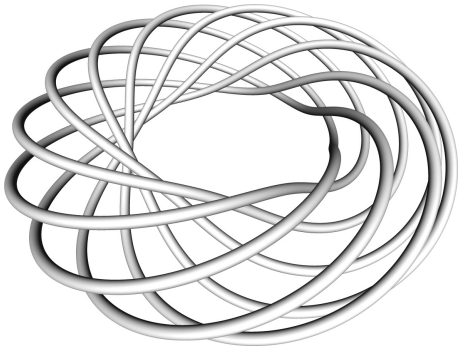

Fig. 2. Toroidal trajectory. The figure shows an idealized toroidal trajectory in which the DNA molecule is represented by a tube and whose axis lies in the surface of a toroid.

using computational models (Petrov *et al.*, 2007). In these studies it is suggested the toroidal shape may be in part a consequence of the elongated shape of some capsids.

1.3. *Spiral-Fold Model*

This model was initially proposed by Earnshaw and Harrison in 1977 and was later supported by eroding experiments on T4 phages (Black *et al.*, 1985). In these experiments partially packed DNA molecules were eroded using ion etching techniques. The illustration in Fig. 3 shows an idealized image of a spiral-fold model. DNA strands run parallel to each other making sharp turns due to the interaction between the genome and the viral capsid. In elongated capsids and in order to avoid unnecessary bending the DNA remains straight for as long as possible. This forces the DNA to eventually take a sharp turn. Current molecular dynamics simulations show configurations that are more consistent with other models. However, sharp turns are also observed in these models (Petrov *et al.*, 2007).

1.4. *Liquid Crystalline Model*

Extreme concentrations of DNA fibers form liquid crystalline phases (Strzelecka *et al.*, 1988; Sikorav *et al.*, 1994). In a liquid crystal state the DNA fibers are locally organized in stacks of parallel planes (with possibly some twisting between the planes) but there is no long-range order between the stacks or the fibers in the stacks. Lepault and colleagues

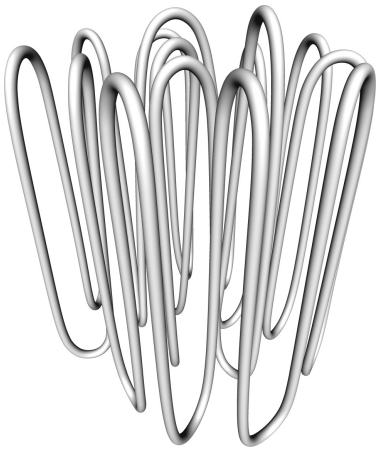

Fig. 3. Spiral-fold model. Black *et al.* proposed that in a T4 capsid the DNA would fold into 700 segments parallel to each other containing between 160 bp and 300 bp forming a total of about 15 shells.

(Lepault *et al.*, 1987) proposed this model after analyzing vitrified samples of phages Giant T4, T3 and Lambda phages by X ray diffraction and direct imaging. Similar findings have been previously reported for herpesviruses (Booy *et al.*, 1991) and also in a later report for DNA in lipid capsules used for gene therapy (Schmutz *et al.*, 1999). Figure 4 shows an idealized image of this model where only the organized fibers are shown.

1.5. *Other Models*

Ball of string model. This model was described in (Richards, 1973) as "a sphere generated by winding the DNA in series of concentric circles, each being approximately a great circle, with the planes containing successive circles oriented at random" and was quickly rejected by experimental evidence (Earnshaw and Harrison, 1977).

Random model. Due to repulsion between DNA fibers and the extreme bending of the DNA inside the capsid it is possible that some of the features described above could be obtained by simply randomly packing the DNA chain in the capsid. We will review some of the properties of the random model in detail. Investigating a random model is useful because it provides a reference for random fluctuations in the packing models described

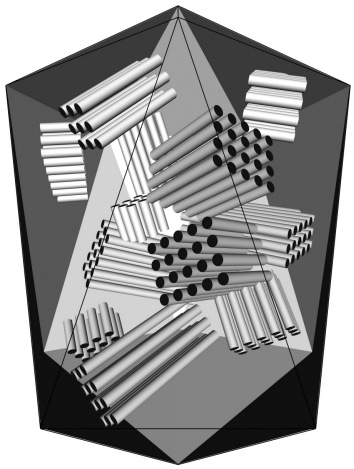

Fig. 4. Liquid crystalline model. In liquid crystalline models, DNA fibers are organized locally but not globally. Sharp turns of the DNA, believed to characteristic of liquid crystalline phases of long molecules, occur in the space between domains. This liquid crystalline phase is a consequence of the extreme density and pressure of the DNA inside the capsid. It is important to highlight that such a model would not provide information on the DNA packaging reaction.

above. An illustration of random packing is shown in Fig. 5. The essential feature of a random model is that the points that define the trajectory of the DNA molecule follow a specific random distribution. This distribution is usually difficult to obtain analytically except for very simple cases (Millet, 1993; Arsuaga *et al.*, 2007). Therefore researchers are performing computer simulations to approach these distributions.

In this review we introduce a new approach to investigating chromosome organization in bacteriophage capsids. Our studies are based on experimental work initiated by Liu and colleagues on P2 and P4 bacteriophages. In Liu *et al.*, 1981a and 1981b it was found that DNA molecules extracted from P2 and P4 tailless mutants were mostly knotted circles. Further studies (Wolfson *et al.*, 1983) found that knots are also observed in wild type P4 phages and that the number of knotted circles (i.e. knotting probability) depends on the length of the DNA being packed. In this review we will argue that these molecules become cyclic inside the capsid, or soon after the disruption of the capsid, and therefore can be used ofto retrieve information about the DNA organization inside the viral capsid. Furthermore we will explain how the DNA knot spectrum can provide

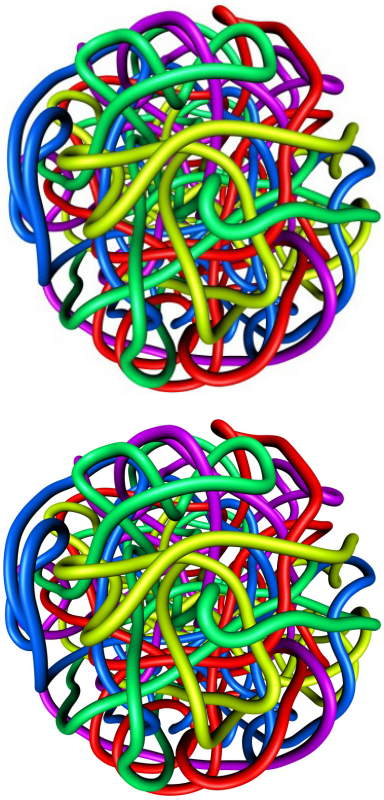

Fig. 5. Random Packing. In the case of random packing the vertices along the trajectory follow a distribution of points in space that is usually difficult to express analytically.

new insights into the spatial organization and geometry of DNA inside the viral capsid.

2. Knot Theory and Its Applications to Molecular Biology

2.1. *Mathematical Knots*

In this section a brief introduction to mathematical knot theory is given. For a more detailed exposition of the basic concepts in Knot Theory we refer the reader to standard texts (Adams, 2004; Murasugui, 2007).

A *knot* K is a simple (i.e. without self-intersections) closed curve in R^3. One studies knots by projecting them into a plane; a projection in which each crossing involves two strings in transverse intersection is called a *regular projection*. If the over/under information for the strands at each intersection in a regular projection is kept, represented by a break in the undercrossing strand, then the projection is called a *knot diagram*. Figure 6 shows some examples of knot diagrams.

Two knots K_1 and K_2 are *equivalent* if it is possible to elastically move K_1 into K_2 without introducing self-crossings. Equivalent knots are representatives of the same *knot type*. Two knots K_1 and K_2 are equivalent if and only if any regular diagram for K_1 can be converted to a regular diagram for K_2 via a sequence of *Reidemeister moves*, as illustrated in Fig. 5.

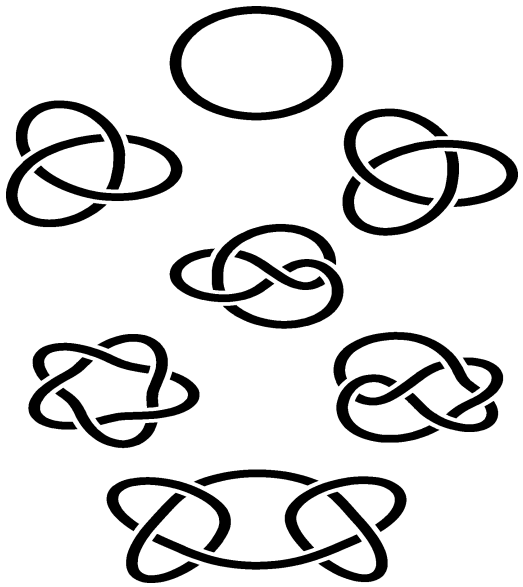

Fig. 6. Examples of knots. The figure shows (from top to bottom) the trivial knot (unknot), the positive (left, denoted here by 3_1) and negative (right, referred here by the mirror image of 3_1 or 3_{1*}) trefoil knots with three crossings each, the four crossing figure-eight knot (center- denoted here by 4_1), the five crossing negative twist knot (right, denoted here 5_{2*}) and the positive torus knot (left, denoted here by 5_1). The knot on the bottom is a six crossing composite knot composed of two trefoils of opposite sign (square knot). From this collection of knots, the trivial knot, the four and six crossing knots have mirror images that are equivalent (achiral), whilst the rest are all chiral. For the trefoil knots, each member of the chiral pair is shown.

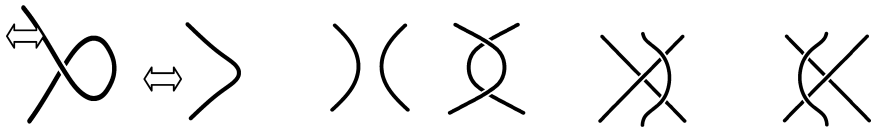

Fig. 7. The three Reidemeister moves. Performing any of these moves in a knot diagram is equivalent to elastically deforming the knot. These moves are used to simplify the projection of the knot without changing the knot type, and therefore facilitate the identification of the knot.

Given two knots, one can construct another knot by cutting a small segment from each of the knots and pasting their endpoints together. The resulting knot is called a *composite knot*. For instance in Fig. 6 the square knot is a composite knot made of two trefoils of opposite sign. Knots that are not composite are called *prime* knots. In Fig. 6, all of the knots except the trivial knot and square knot are prime knots.

The minimum number of crossings among all possible knot diagrams for K is called the *crossing number* of K. A knot diagram that realizes that minimum crossing number for K is called a minimal diagram for K. Figure 6 shows minimal diagrams. A concept that has proven to be very useful in biology is the *average crossing number*. It is given by the average number of crossings over all possible regular projections of K(as a fixed space curve). A second important concept is that of *chirality*. A knot K is said to be *achiral* if it is topologically equivalent to its mirror image, otherwise it is said to be *chiral*. In order to obtain the mirror image of a knot, take any diagram for the knot, and reverse all the crossings. In Fig. 6, the right-handed trefoil is topologically distinct from its mirror image, the left-handed trefoil. The figure-eight knot and square knots are achiral since one can find a sequence of Reidemeister moves that can transform the diagrams shown to their mirror images. A geometrical concept that helps quantify the chirality of a knot is the *writhe*. Given a knot K (as a fixed curve in space) and an orientation on the knot (i.e. a direction of travel along the knotted chain) one can define the *projected writhe* of the knot by assigning an orientation to the knot, and at each crossing in a regular diagram of K either a $+1$ or a -1 is assigned following the right hand rule crossing sign convention (Fig. 8) for oriented skew lines in space. Note that the crossing signs are independent of knot orientation, because changing the orientation of the knot changes both of the arrows at a crossing, leaving the crossing sign unchanged.

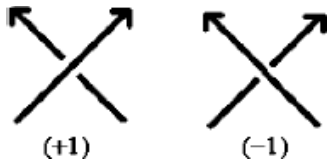

Fig. 8. Crossing Sign Convention.

If one averages the projected writhe over all projections of a fixed spatial curve K, one obtains the *writhe* of the knot K. The writhe is a measure of non-planarity of the spatial knot. For the knot diagrams in Fig. 6, the projected writhe of the unknot, figure-eight knot and square knots are each zero; the projected writhe of the positive trefoil is $+3$, the projected writhe of the negative trefoil is -3, the projected writhe of the negative twist knot is -5, and the projected writhe of the positive torus knot is $+5$. If a knot K is achiral the average of the writhe of a random sample of curves of fixed knot type K tends to zero as the number of knots in the sample goes to infinity.

The computational identification (classification) of knots is done by using algebraic knot invariants that can be calculated from any regular projection of the knot. If a given invariant is different for two knots, then the knots are different. If a given invariant is the same for two knots, the knots may or may not be different; that invariant fails to distinguish them. In computational studies, polynomial invariants such as the Alexander, Jones and HOMFLY polynomials are commonly used. The Alexander polynomial, denoted by $\Delta(t)$, is easier and faster to compute than the others, but it is not as powerful as the HOMFLY or the Jones polynomials in classifying knots. An important drawback of the Alexander polynomial is that it cannot detect knot chirality. It is known that the computation of the Jones polynomial is NP-hard (i.e. cannot be computed in time, which is polynomial in the number of knot crossings) (Jaeger *et al.*, 1990). In order to make computations more efficient, the Alexander polynomial evaluated at $t = -1$ (denoted by $\Delta(-1)$) is commonly used. However theoretical studies have shown that there exist many non-trivial knots with trivial Alexander polynomials ($\Delta(t) = 1$) and it is not clear that $\Delta(-1)$ can be used to estimate knot probability for arbitrarily knot ensembles (Garoufalidis and Teichner, 2004). Another approach is to use look-up tables of algebraic descriptors of knots called Dowker codes (Hoste and Thistlethwaite, 1999; Micheletti *et al.*, 2006). The disadvantage of look — up tables is that the exact tabulation of knots excludes knots of more than 16 crossings.

2.2. *Knots and DNA*

Knots and links in circular DNA molecules are of biological interest because they can detect and preserve information about DNA folding and about enzymes that manipulate DNA. DNA knots can be obtained in the test tube (*in vitro*) as products of biochemical reactions or can be designed as building blocks for nanotechnology purposes (e.g. Mueller *et al.*, 1991; Seeman 2003). Figure 9 shows a DNA trefoil knot produced *in vitro* by reacting an unknotted nicked duplex DNA plasmid with bacterial topoisomerase I.

DNA knots obtained as the product of biochemical reactions help analyze the process under study. For instance, knots formed by random cyclization of linear molecules have been used to quantify the effective diameter of the DNA (Rybenkov *et al.*, 1993; Shaw and Wang, 1993) and the effects of confinement on the cyclization of DNA molecules (Arsuaga *et al.*, 2002). DNA knots have been also key to unveiling the mechanism of site-specific recombination enzymes (e.g. Buck and Flapan, 2007; Darcy *et al.*, 2006; Ernst and Sumners, 1990; Wasserman and Cozzarelli, 1986; Grainge *et al.*, 2007; Vazquez and Sumners, 2004; Vazquez *et al.*, 2005; Vetcher *et al.*, 2006), the structure of the Mu transpososome (Darcy *et al.*, 2007; Pathania *et al.*, 2002), the mechanism of type-2 topoisomerases (Flammini *et al.*, 2004; Hua *et al.*, 2001; Liu *et al.*, 2006; Roca, 2001; Rybenkov *et al.*, 1997; Vologodskii *et al.*, 2001), the mechanism

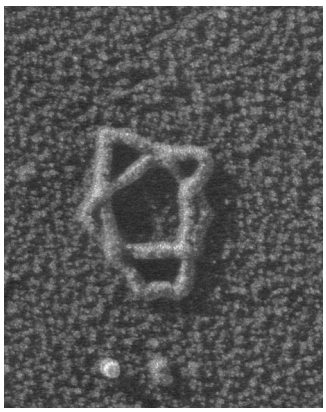

Fig. 9. Electron micrograph of a +DNA five crossing torus knot obtained after the action of a site-specific recombinase on a dsDNA plasmid. (Figure kindly provided by Nancy Crisona.)

of condensins (e.g. Kimura *et al.*, 1999), and the structure of replication bubbles (e.g. Olavarrieta *et al.*, 2002). In most of these cases, the development of mathematical and computational tools has helped greatly in the analysis of the experimental results. Yet, the origin of some DNA knots found in other biological systems remains mostly unknown. Examples include *Escherichia coli* cells harboring mutations in the GyrB or GyrA genes (Shishido, 1987), and the cauliflower mosaic virus (Menissier *et al.*, 1983). A recent *in vivo* study (Diebler *et al.*, 2007) shows that knotting of a DNA plasmid that contains a gene for drug resistance leads to an increased mutation rate of the bacterial host, inhibition of replication and cell death when the knot prevents expression of the drug resistance gene.

Knotted and linked (catenated) DNA molecules are stable in the laboratory and can be characterized experimentally by gel electrophoresis and electron microscopy. The simplest method to detect DNA knots in the laboratory, and the one used in the experiments described below, is agarose gel electrophoresis. In this assay, DNA molecules are run at low voltage and knots migrate approximately according to their average crossing number (Vologodskii *et al.*, 1998). Fig. 9 shows a DNA five crossing torus knot produced *in vitro* by reacting an unknotted DNA plasmid with a site-specific recombinase.

A second method to identify knots in the laboratory is by microscopy, using both transmission electron and atomic force microscopes. Since we have not used these assays in our studies we will not provide details here and instead refer the reader to Zechiedrich and Crisona, 1999. However it is important to highlight that the advantage of microscopy over gel electrophoresis is that it allows detection of DNA knot chirality. For example, in Fig. 9, the + trefoil DNA knot has been coated with the RecA protein to enhance resolution of DNA crossings in the electron micrograph.

2.3. *Computer Models of DNA Knotting*

Both Molecular Dynamics (MD) and Monte-Carlo (MC) methods have been used to study circular DNA in solution and in confinement. Here we will review only those studies where Monte-Carlo methods were used.

Monte-Carlo simulations of off-lattice (continuum) and on-lattice models have been used to generate ensembles of simulated closed chains. Initial computational efforts as well as some more recent studies have

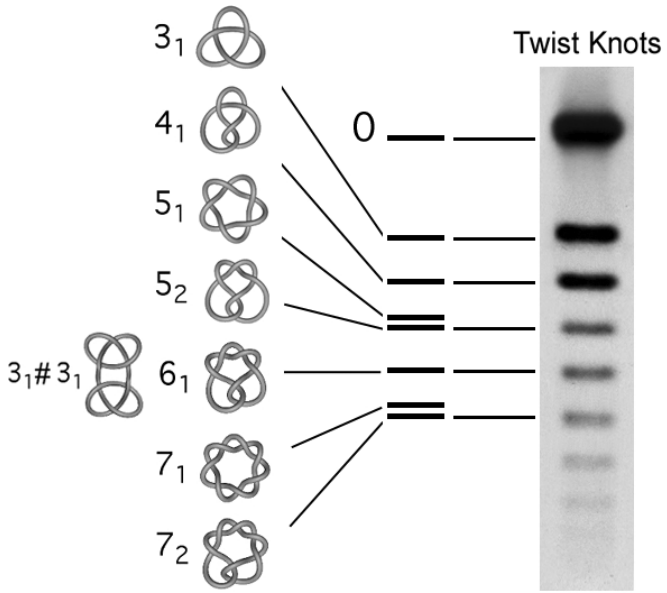

Fig. 10. Electrophoretic velocity of DNA knots. Figure adapted from Arsuaga *et al.*, 2005. DNA knots migrate approximately proportional to their average crossing number. The figure shows the migration pattern of twist knots generated by the action of type II topoisomerase from bacteriophage T4 observed first in Wasserman and Cozzarelli, 1991.

used different lattices such as the tetrahedral lattice (Frank-Kamenetskii *et al.*, 1975) or the simple cubic lattice (Tesi *et al.*, 1994; Hua *et al.*, 2005). However off-lattice models are usually preferred to study DNA properties. These models include the equilateral random polygon model (ERP) (also known as the freely-jointed chain model), the Gaussian random polygon model (GRP), and the wormlike chain model (Cantor and Schimmel, 1980). Among the off-lattice models, the worm-like chain is believed to be the one that best approximates DNA in dilute solution. For the scales considered in the work reviewed here and for comparison with bulk experiments both the wormlike chain and the ERP give very similar results. One advantage of the ERP is that a number of properties, such as the knotting probability and average crossing number, have been characterized not only computationally but also analytically. These results provide a solid foundation for the computational studies (e.g. Diao *et al.*, 1994; Diao 1995).

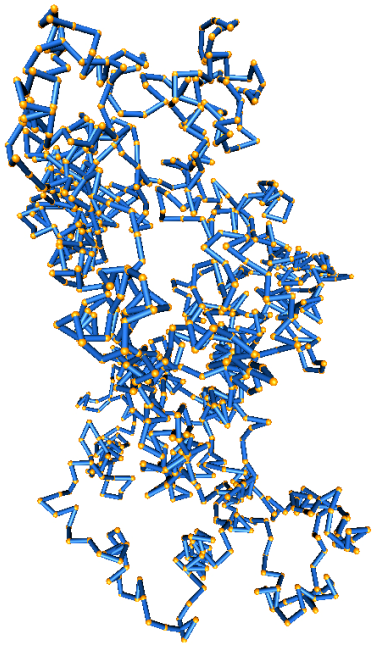

Fig. 11. An equilateral random polygon. An equilateral random walk with 1000 independent segments is shown. The equilateral walk was generated using the hedgehog algorithm and rendered using Knotplot (Scharein, 1998).

2.4. *Algorithms for Generating Ensembles of Random Knots*

There are a number of algorithms that generate *equilateral random polygons*. These include the crankshaft (Klenin *et al.*, 1988), the hedgehog (Klenin *et al.*, 1988), and the generalized hedgehog algorithm (Varela *et al.*, 2009). The crankshaft algorithm generates a Markov chain in the space of all ERPs (or wormlike polygons) of fixed length. This Markov chain is known to be ergodic (Millet, 2000). However this algorithm is computationally very inefficient due to the high correlation between consecutive samples. The hedgehog algorithm on the other hand is more efficient than the crankshaft; however it is unknown whether it is ergodic or not. The generalized hedgehog is the only algorithm known to be ergodic and fast.

Once a polygon has been generated by any of the methods described above, a projection is taken and the knot type computed. This process

however may be very costly because the knot projection may have many extraneous crossings and the computation of any knot invariant depends on the number of crossings. In order to lower the computational burden one can smooth the polygon before computing its projection (e.g. Micheletti *et al.*, 2006). This smoothing is done by perturbing the knot across those triangles that are spanned by adjacent edges and that do not intersect other segments of the knot. For instance, in Micheletti *et al.*, 2006, it was observed that unknots of length $N = 400$ edges confined to a sphere of $R = 10$ typically have approximately 400 crossings! By applying these simplification methods, polygons were simplified to lengths of $N' = 30$ with approximately 30 crossings. Next a projection of the knot is taken and Reidemeister moves (Fig. 7) performed to further remove extraneous crossings. Despite these efforts the knot diagrams may still have many extraneous crossings. In order to determine if a configuration is knotted or not, it is necessary to compute an algebraic invariant. A quick and reasonably efficient determination of curve knotting is computing $\Delta(-1)$. Computing the knot type is a more challenging problem and researchers have either computed the Alexander polynomial $\Delta(t)$ evaluated at different values of t (e.g. $t = -2, -3$, see Arsuaga *et al.*, 2005; Mansfield, 1994) or have used more sophisticated methods such as the HOMFLY polynomial (Millet, 1994; Micheletti *et al.*, 2006; 2008).

3. DNA Knots in Bacteriophage P4

3.1. *Bacteriophage P4*

Bacteriophage P4 is a satellite of bacteriophage P2. The P4 capsid has icosahedral symmetry ($T = 4$) (Dockland *et al.*, 1992) and its genome is an 11.7 kb linear double-stranded DNA molecule with two 16 bp long cohesive ends (Wang *et al.*, 1973) called *cos* sites. In the mature phage capsid, one of the DNA ends is attached near the connector region (Chattorraj and Inman, 1974), while the position of the other end is unknown. The attachment of one of the ends to the entry region prevents the two *cos* ends from annealing to each other and forming a circle inside the capsid. It is likely that this anchoring facilitates the DNA in leaving the capsid as a linear molecule through the tail at the time of infection.

However, most DNA molecules extracted from bacteriophage P4, as detailed in Section III.2, are non-covalently closed circular molecules and

have a very high percentage of nontrivial and highly complex knots (∼50%) (Arsuaga *et al.*, 2002; Wolfson *et al.*, 1985). The percentage of knots is higher (∼95%) for DNA extracted from tailless mutants of the bacteriophage P4 (Arsuaga *et al.*, 2002; Liu *et al.*, 1981a; Liu *et al.*, 1981b; Wolfson *et al.*, 1985). Interestingly, the percentage of knots is also higher in P4 strains with shorter genomes (∼10 kb) than in those strains with full genomes (Wolfson *et al.*, 1985).

What is the origin of these knots? In order to form a knot, both ends of a linear DNA chain need to meet and anneal through their *cos* ends. This cyclization reaction results in non-covalently closed circles (since the DNA backbones remain unsealed). The cyclization reactions of purified linear P4 DNA molecules in free solution have been extensively studied. In these experiments, DNA molecules were first extracted and linearized by thermal denaturation of their *cos* ends. Upon cyclization in dilute free solution and at physiologic ionic conditions, it was found that the knotting probability for P4 DNA molecules is about 3% and that most of the knots (> 90%) had only three crossings (Rybenkov *et al.*, 1993). Therefore DNA knots have a non-zero probability of occurrence during any cyclization reaction. We ask then why the knotting probability and complexity are so high in molecules extracted from P4 phages. Another important question is why this probability reaches saturation levels in tailless mutants or in the genome deletion strains.

3.2. *Experimental Methods: P4 Production and DNA Extraction*

Bacteriophage P4 vir1 del22, which carries a 1.7 Kb deletion compared with the wild type P4, are used in the laboratory as source of knotted DNA. These phages with a shorter DNA (∼10 Kb) show higher DNA knotting percentages than those with full genomes (Wolfson *et al.*, 1985). In order to obtain a maximum yield of knotted molecules, P4 vir1 del22 is first amplified in an *E. coli* strain (for example, C-1895) lysogenic for the helper prophage P2. Stocks of P4 vir1 del22 are then used to infect an *E. coli* strain lysogenic for P2 (for example, C-8001) that carries the mutation amH13 in the H gene. Because gene H encodes part of the P4 phage tail, such strain produces then P4 vir1 del22 particles without tail (tailless mutants).

After bacterial lysis, bacteriophages are collected as described in Isaksen *et al.*, 1999. Briefly, phages are precipitated with polyethylene glycol 8000 and then dissolved in phage buffer (10 mM MgCl2/10 mM Tris·HCl, pH 7.2/130 mM ammonium acetate). Phage particles can then be purified by equilibrium density centrifugation across a cesium chloride gradient (14 h at 45,000 rpm in an NVT65 Beckman rotor). Banded viral particles are dialyzed against phage buffer and their DNA is extracted with phenol and phenol/chloroform. The viral DNA is then dialyzed against TEN buffer (10 mM Tris·HCl, pH 8/1 mM EDTA/100 mM NaCl) and kept at 4°C.

An optimal way to qualitatively and quantitatively analyze knotted DNA is by means of a high-resolution two-dimensional gel electrophoresis (Trigueros *et al.*, 2001). In this technique, the first gel dimension is run at low voltage, and DNA knots migrate according to their compactness. The second gel dimension is run at high voltage, and DNA knots migrate according to other physical parameters, such as shape and flexibility. In comparison to one-dimensional gel electrophoresis, this procedure segregates the knotted DNA molecules from other unknotted forms of DNA, and partially resolves populations of knots that have the same number of crossings. For the analysis of P4 DNA knots (about 10 Kb size), these gels are prepared as follows. Gel slabs of 0.4% agarose concentration are equilibrated with TBE buffer (100 mM Tris-borate, pH 8.3/2 mM EDTA). In the first dimension, DNA samples run at 0.8 V/cm for 40 h at room temperature. After a 90° rotation of the gel, a second dimension is run in the same electrophoresis buffer at 3.4 V/cm for 4 h at room temperature. After ethidium bromide staining of DNA and photography, gels can be blotted to nylon membranes, and DNA bands radioprobed for phosphorimaging quantification (Fig. 12).

3.3. *Experimental Results*

The quantitative analyses of DNA species resolved by high-resolution two-dimensional gel electrophoresis, like that seen in Fig. 12, indicated the following:

— The 10-kb DNA from the tailless mutant of phage P4 vir1 del22 produces 95% knotted molecules (*i.e.* a knotting probability of 0.95).

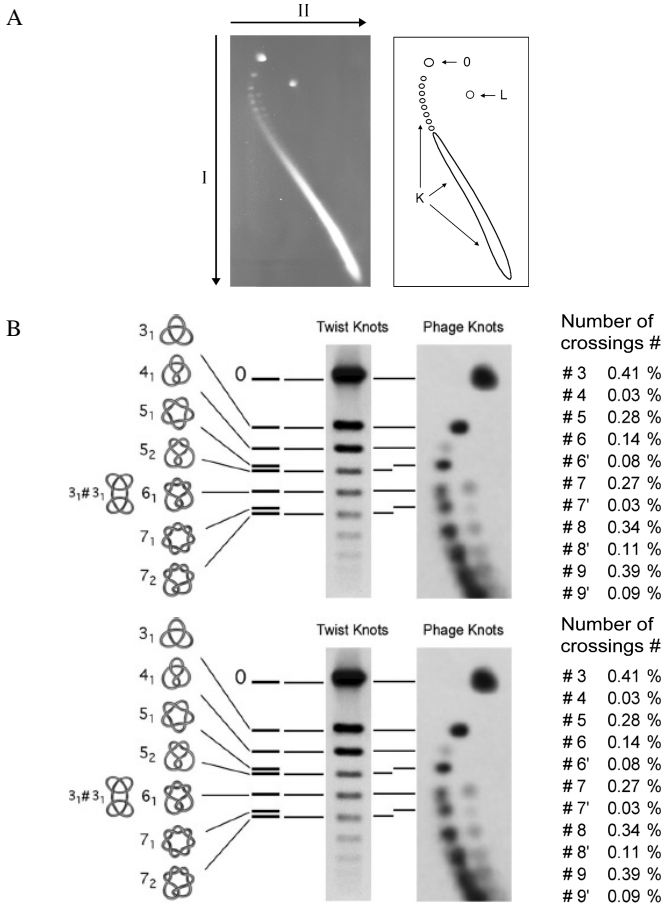

Fig. 12. Qualitative and quantitative analyses of DNA knots resolved by gel electrophoresis (from Arsuaga *et al.*, 2005). A) Knots are identified by 2D gel electrophoresis (data from Arsuaga *et al.*, 2005). In the first dimension (vertically) knots were run for 40 hours at 25 V and were run parallel to a control of twist knots (Wasserman and Cozzarelli, 1991). Under these conditions knots migrate proportionally to their average crossing number. The second dimension (horizontally) is run for under 2 hours at 100 V. Knots migrate differently creating the characteristic bell shape. For details see Trigueros *et al.*, 2001. B) The gel-blot region of less complex knot populations (knots of 3 to 9 crossings) is shown in detail. Gel positions of DNA bands are compared to a marker ladder for twist-type knots that was generated by incubating a supercoiled 10-kb plasmid with equimolar amounts of T4 topoisomerase II (Wasserman and Cozzarelli, 1991). After the knotting reaction, supercoils were removed from DNA by limited nicking in a reaction containing 50 mM Tris (pH 7.5), 5 mM MgC12, 100 pg/ml each DNA and ethidium bromide, and 2 pg/ml DNase I.

— The average estimated number of knot crossings was about 26.7, and the largest estimated number of knot crossings detected was about 40.
— Of the total amount of knotted molecules extracted from P4 tailless capsids, only 2.2% had crossing numbers between three and nine. These populations are of main interest since they can be individually quantified.
— Some of the bands observed in the 1D gel split into a secondary arch of faster velocity in the second gel dimension (denoted as 6–9 and 6'–9' in Fig. 12).
— Densitometer readings of individual knot populations of three to nine crossings revealed that knots of four crossings are severely reduced relative to the other knot populations.
— At low voltage, torus knots (such as 5_1 and 7_1) migrate slightly slower than their corresponding twist knots (5_2 and 7_2) (Vologodskii *et al.*, 1988); this knowledge in conjunction with a marker ladder for twist knots (Fig. 10 and Wasserman and Cozzarelli, 1991) allows the identification of several gel bands of the phage DNA matching the migration of known knot types.
— In the main arch of the gel, in addition to the unambiguous knots 3_1 and 4_1, the knot population of five crossings matched the migration of the torus knot 5_1. The other possible five-crossing knot, the twist knot 5_2 that migrates between and is equidistant to the four- and six-crossing knot populations, appeared to be negligible or absent.
— The knot population of seven crossings matched the migration of the torus knot 7_1 rather than the twist knot 7_2, which has slightly higher gel velocity. Although this visible gel band is likely populated by seven-crossing torus knots, other knot types of seven crossings cannot be excluded.
— There is an evident shortage of the knot subpopulation of seven crossings in the second arch of the gel (denoted by 7' in Fig. 12).

When comparing these knot distributions with those obtained in free solution, we conclude that the DNA particles become cyclic inside the viral capsid.

4. Models of Random Packing

In this section we consider the computational aspects of the problem of DNA knotting in bacteriophages. As discussed above we mainly work

with the equilateral random polygon (ERP) which in this case is confined inside a spherical volume that represents the viral capsid. Before we go in a detailed discussion of the models and the results, a few assumptions need to be highlighted.

The first assumption deals with the flexibility of the chain. The flexibility of the dsDNA chain is given by the persistence length which is 50 nm in free solution. However the radius of the P4 capsid is 45 nm. This strong confinement indicates that there is an effective persistence length inside the capsid that is different from that in solution. Due to these restrictions researchers are forced to make approximations (Arsuaga *et al.*, 2002; LaMarque *et al.*, 2004; Arsuaga *et al.*, 2005; Arsuaga *et al.*, 2007; Micheletti, 2008). In our studies and those of others, the flexibility of the chain is considered by expressing capsid radius R as a multiple of the effective persistence length. The second assumption has to do with the generation of random samples of closed curves inside a sphere of fixed radius. The sample of knots obtained by randomizing polygons in confinement, using the algorithms described above, is believed to be comparable to the sampling process that the DNA undergoes during random cyclization. This assumption has been confirmed when analyzing DNA cyclization experiments in free solution (Rybenkov *et al.*, 1993).

4.1. *Computational Methods*

Modeling DNA under spatial constraints remains a challenge and many of the techniques explained above (Section II.3) become very expensive computationally. The problem of knotting in confined volumes was first addressed by Michels and Wiegel (Michels and Wiegel, 1986) using molecular dynamics simulations. Today, the most common method to generate statistical ensembles of polygons confined to spheres is Monte-Carlo. The simplest approximation is to generate polygons by one of the methods described earlier and reject those that are outside the confining sphere. However when one tries to sample polygons in confinement one finds that the vast majority of rings will be "swollen" and accumulating reliable statistics for highly confined conformations leads to impractically long computing times. A more efficient computational alternative is to generate a succession of conformations chosen using importance sampling criteria (Arsuaga *et al.*, 2002). However this method is still not sufficient for long polygons confined to small volumes. In Micheletti *et al.*, 2006 and

2008 a dramatic improvement on these calculations was accomplished. First, calculations were done in the conjugated ensemble rather than in the ensemble of a sphere of fixed radius. Second, a faster sampling along the Markov Chain was achieved by a multiple Markov chain scheme, in which all "replicas" of the system were run in parallel, running each chain at a different temperature. Third, the data were collected and analyzed using histogram reweighting techniques.

A second approach consists of initially defining points that are physically within the confining region. Both continuous and discrete models have been implemented (Arsuaga *et al.*, 2007; Mansfield, 1994; Millet, 2000). Current continuum models are known to be non-realistic for simulating DNA since the flexibility of the chain is not well defined. However these studies have shown that, qualitatively speaking, results are similar to those obtained by other methods. The advantage of these methods is that they are drastically more efficient and also more amenable for the development of analytical results than the methods described above.

4.2. *Results*

4.2.1. *Knotting probability, average crossing number and writhe in confined random polygons*

In Arsuaga *et al.*, 2002, 2005 and in Micheletti *et al.*, 2006 the effects of confinement on the ERPs knotting probability were analyzed by Markov-chain Monte Carlo simulations as explained above. Figure 13 (from Micheletti *et al.*, 2006) shows the knotting probability of closed ideal chains confined to spherical volumes of various radii (in multiples of edge length) as a function of the total chain length (number of edges in the equilateral polygon). As one would expect, and in agreement with previous models, the knotting probability decreases with increasing radii and decreasing polygon length (since the effects of the confining sphere are gradually removed under these two conditions). In Micheletti *et al.* (2006), it was found that the ratio of probabilities for polygons confined with respect to those that are not confined is given by $P(N,1/R)/P(N,$ no confinement$) \sim \exp(N^a/R^3)$ where $a = 2.15$.

In addition, the complexity of the knots was estimated in Arsuaga *et al.*, 2002 by computing the average number of crossings of the knot. These results are summarized in Fig. 14. As expected, it was observed that for

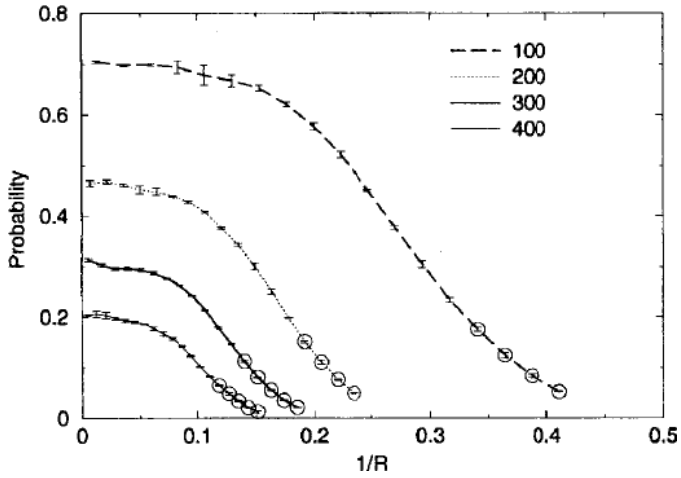

Fig. 13. Knotting probability for chains of length N with $50 \leq N \leq 450$ as a function of $1/R$. Open symbols denote the region where the fraction of unknown knots is between 10% and 50%. No data is plotted where the fraction of unknown knots exceeds 50% (Figure Adapted from Micheletti *et al.*, 2006).

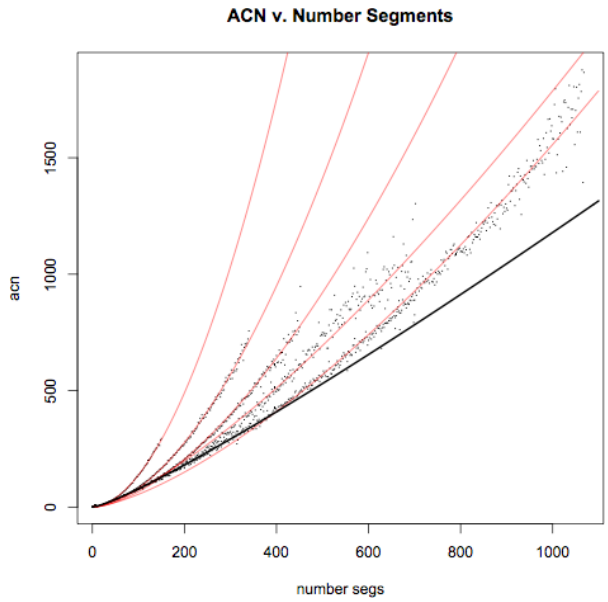

Fig. 14. Average Crossing Number of ERPs in Confined Volumes. The graph shows results for chains of length N with $20 \leq N \leq 1000$ confined to radii ($R = 2, 3, 4, 5, 7$) Results for non confinement are shown by the black line.

a fixed confining volume the complexity increased with the length of the polygon. For fixed length, the complexity decreased when the confining radius was increased.

In Micheletti *et al.* (2006), a preliminary study of the average writhe of polygons confined to spheres was performed. Since the average value of the writhe is zero one needs to compute the absolute value of the writhe or the square of the writhe to estimate its variation with the confining volume. It has been rigorously shown that the absolute values of the writhe of polygons in the simple cubic lattice without confinement increase as N^a with $a = 0.52 \pm 0.04$. In Micheletti *et al.*, 2006 it was found that the value of the exponent a changes when in confinement.

4.3. *Distributions of Knots for Low Crossing Numbers: Non-Random Packing of DNA in Bacteriophage P4*

In Arsuaga *et al.* (2005) and in Micheletti *et al.* (2006) and 2008 the distributions of knots with low crossing number, that is knots with fewer than 6 crossings, were computed and plotted as functions of the inverse of the radius or of the length of the polygon.

Figure 15 (left) shows the R dependence of the probability to observe a trefoil knot $P(3_1)$ for various lengths of the polymer ring. As for the unknotting probability there is a range of R ($R > Rc$) for which $P(3_1)$ does not change too much with R. This "plateau" is more visible for small values of N when, for sufficiently small confining radii, $P(3_1)$ is a monotonic-nonmonotonic curve with one maximum. As the confinement

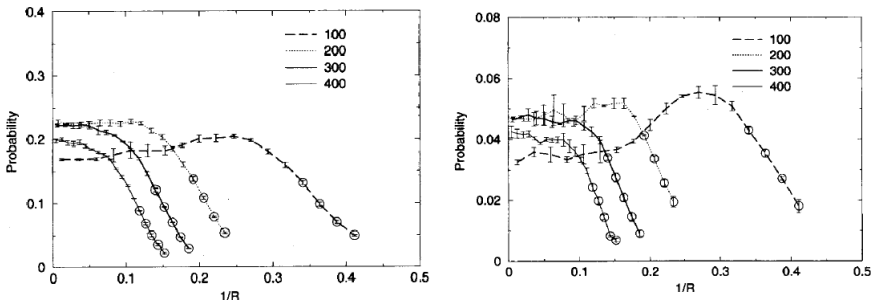

Fig. 15. Knot probabilities as a function of N and $1/R$. The plots are for the 3_1 and 4_1 knots.(Figure adapted from Micheletti *et al.*, 2006)

radius R is further reduced, $P(3_1)$ decreases in favor of more compact conformations. For longer polymers ($N > 125$) the maximum becomes progressively less evident and we observe a shoulder for small values of $1/R$ that disappears for $N = 400$.

Figure 15 (right) show plots analogous to the one in Fig. 13 but now refer to the probability of forming four crossing knots. The trend observed for the trefoil knot is also observed for the four crossing knot and also other prime knots with less than 7 crossings. Indeed, all the curves for short polymers (small N) have a maximum as a function of the inverse radius while those for longer polymers decrease monotonically with $1/R$.

In Arsuaga *et al.* (2005), a comparison between the distributions of knots for the knot families $3_1, 4_1, 5_1$ and 5_2 was made for different lengths and radii. It was found that in all tested cases the probability of the four crossing knot was higher than that of each of the fives (but not combined). Furthermore, as it is the case for knots without confinement, the twist five (5_2) was more probable than the toroidal five (5_1). This result has been observed for different polymer models in confined volumes such as the Uniform Random Polygon (Millet) and the wormlike chain (Micheletti *et al.*, 2008). Importantly, Fig. 16 shows the simulated distribution of knots that differs strongly from the one observed experimentally (see Fig. 12). Since the simulated case represents the distribution of knots obtained for random embeddings of closed curves in confined volumes it was concluded that the packing in bacteriophage capsids cannot be random (see Fig. 16).

4.4. *Knotting Probability and Complexity in Biased Models*

How can we explain the differences between the knot distributions obtained from random embeddings and those observed experimentally?

Chiral models

In Arsuaga *et al.*, 2005 it was argued that elevated writhe was responsible for these differences in the knot distribution. One interesting observation is that the 4_1 knot is achiral, that is it is equivalent to its mirror image. This property is somehow captured by the writhe of the molecule since achiral knots have writhe 0. In fact in (Rensburg *et al.*, 1993) it was found that random polygonal realizations of the 4_1 knot in free space

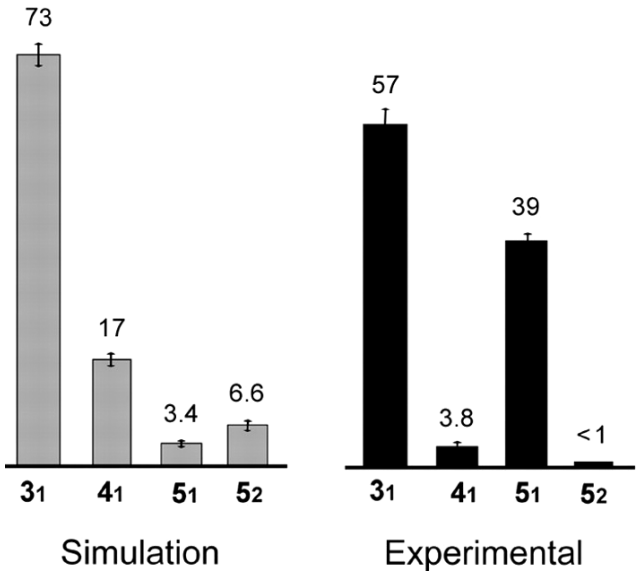

Fig. 16. DNA packing in bacteriophage P4 is not random. Comparison of the computed probabilities of the knots 3_1, 4_1, 5_1, and 5_2 (for polymers of length $n = 90$ randomly embedded into a sphere of radius $R = 4$) with the experimental distribution of knots. The relative amount of each knot type is plotted. Note that the fractions of knots 3_1 and 4_1 plausibly formed in free solution are not subtracted from the experimental distribution. If these corrections are considered, the relative amount of knot 4_1 is further reduced. (Figure adapted from Arsuaga *et al.*, 2005)

produce a family of polygons whose writhe distribution for any polygonal length is a Gaussian curve with zero mean and whose variance grows as the square root of the length. As discussed above the confinement alone increases the writhe (Micheletti *et al.*, 2006). In order to obtain conformations with elevated writhe, Arsuaga and colleagues sampled polygons whose writhe was above the average writhe induced by the confinement. It was found that the distribution of knots changes dramatically and produces results that are comparable to those obtained experimentally. Figure 17 shows how the knot distribution for the 4 and 5 crossing knots changes with increasing writhe. In Arsuaga *et al.* it was proposed that the main reason for the scarcity of the knot 4_1 is a writhe bias imposed on the DNA inside the phage capsid. A drop of the probability of the knot 4_1, as well as a rapid increase of the probability of the torus knot 5_1 but not of the twist knot 5_2, readily emerged by increasing the writhe rejection

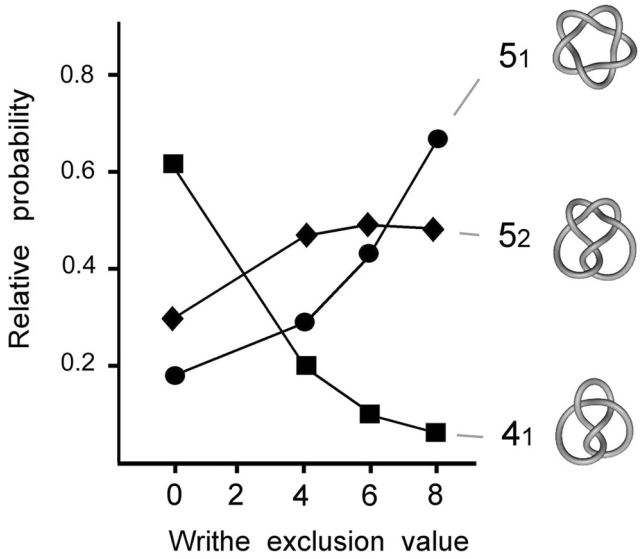

Fig. 17. Effect of writhe-biased sampling on the probability of knots 4_1, 5_1, and 5_2. Figure adapted from Arsuaga *et al.*, 2005. Increasing the value of the writhe modulates the distributions of observed knots. Importantly it decreases the probability of the four crossing knot as well as the five twist knot while at the same time increasing the five toroidal knot.

value. These writhe-induced changes in the knot probability distribution are independent of the number of edges in the equilateral polygon and the sphere radius length. Accordingly, previous studies had shown that the mean writhe value of random conformations of a given knot does not depend on the length of the chain but only on the knot type and that these values are model-independent (Rensburg *et al.*, 1993).

5. Conclusions and Future Challenges

Here we have reviewed some of our topological approaches to the problem of genome organization in bacteriophage P4. There are two main conclusions: First, P4 DNA knots are formed either inside the bacteriophage capsid or soon after its disruption. This observation indicates that the knotting observed is driven by the confinement of the capsid. It also suggests that P4 knots can be used as reporters for DNA packing in P4 phages (Arsuaga *et al.*, 2002). Second, analysis of P4 knot distributions

suggests that the viral genome is chirally organized (Arsuaga *et al.*, 2005). This property had not been previously observed using other experimental techniques. Next we briefly describe some of the latest advances that will require further exploration.

i) *Excess writhe is consistent with the knotting probabilities observed in deletion mutants.*

Results obtained by Wolfson *et al.* that show that shorter DNA chains extracted from bacteriophage P4 have higher knotting probabilities than long ones have long puzzled researchers. A possible explanation was first provided by Blackstone *et al.* (2007), where it was shown that the knotting probability of simulated ERPs with high writhe values decreases for increasing chain length until a minimum is reached. After this point the knotting probabilities increase as expected.

ii) *Currently proposed DNA packing models are not consistent with P4 data.*

Simulations of random closure of the DNA molecules packed as described in Section 1 show very little knotting. This observation suggests that current models for DNA organization inside phage capsids are simplistic and do not completely reflect the complexity of the packing. A new model called "random spooling model" has been proposed by Arsuaga and Diao to address this problem (Arsuaga and Diao, 2008). This model simulates concentric spools in which the DNA strands from different concentric layers are allowed to interweave.

iii) *Dealing with knot complexity.*

As described in the sections above, only about 2% of the knots can be identified using current experimental methods. One way to estimate the complexity of those knots is by creating deletion mutants that can be better characterized by gel electrophoresis or microscopy. A second option is to treat samples with type II topoisomerases. Topoisomerases can unknot knots and therefore by studying possible unknotting pathways one may infer the complexity of the initial population. In Hua *et al.* (2005), a computational study simulating this approach was proposed and equilibrium distributions of knots were found.

iv) New deletion mutant strains may help better identify the knots observed as well as the packing arrangement.

One drawback of this natural system is that knot formation is restricted to the viral DNA. Hence, we asked whether other DNA molecules of

length and sequence different than P4 DNA could be packaged inside P4 capsids and recovered also as highly knotted forms. Accordingly, we envisaged that a bacterial plasmid containing the P4 *cos* sequence (i.e., a P4 cosmid) could be cleaved and threaded into a viral capsid in the course of a bacterial infection by phage P4. Therefore, we constructed different P4 cosmids and introduced them in bacteria lysogenic for P2. These bacteria were then infected with P4 phage and the DNA in newly formed viral particles was analyzed. We found that cosmids as small as 5 kb were packaged inside P4 capsids. More interestingly, as well as P4 DNA, such cosmids were recovered in the form of highly knotted DNA circles (Trigueros and Roca, 2006). These results further indicate that P4 knots are not a consequence of some property of the viral DNA, and also demonstrate that complete filling of the capsid is not essential for the packing of individual DNA molecules. With this biological system, DNA molecules of varying length and sequence will be shaped into very complex and heterogeneous knotted forms. These molecules could be produced in preparative amounts suitable for systematic studies and novel applications.

Acknowledgements

We thank R. Scharein for his help in generating the DNA packing figures shown in Section 1 as well as Figures 6–8. Also we want to thank R. Varela and B. Borgo for their help in generating Figs. 11 and 14 respectively. J. Arsuaga is supported by NIH grant 2S06GM52588–12.

References

1. C. Adams. *The Knot Book: An elementary introduction to the mathematical theory of knots, knots.* Revised reprint of the 1994 original, (American Mathematical Society, Providence, RI, 2004).
2. J. Arsuaga and Y. Diao. DNA knotting in spooling like conformations in bacteriophages. *Journal Computational and Mathematical Methods in Medicine* **9(3–4)**:303–316 (2008).
3. J. Arsuaga, R.K. Tan, M. Vazquez, D.W. Sumners and S.C. Harvey. Investigation of viral DNA packaging using molecular mechanics models. *Biophys. Chem.* **101**:475–484 (2002).
4. J. Arsuaga, M. Vazquez, S. Trigueros, D.W. Sumners and J. Roca. Knotting probability of DNA molecules confined in restricted volumes: DNA knotting in phage capsids. *Proc. Natl. Acad. Sci. USA* **99**:5373–5377 (2002).

5. J. Arsuaga, M. Vazquez, P. McGuirk, S. Trigueros, D.W. Sumners and J. Roca. DNA knots reveal a charal organization of DNA in phage capsids. *Proc. Natl. Acad. Sci. USA* **102**:9165–9169 (2005).
6. J. Arsuaga, T. Blackstone, Y. Diao, E. Karadayi and Y. Saito. Knotting of uniform random polygons in confined spaces. *J. Phys. A: Math Gen.* **40**:11697–11711 (2007).
7. K. Aubrey, S. Casjens and G. Thomas. Secondary structure and interactions of the packaged dsDNA genome of bacteriophage P22 investigated by Raman difference spectroscopy. *Biochemistry* **31**:11835–11842 (1992).
8. L. Black, W. Newcomb, J. Boring and J. Brown. Ion etching bacteriophage T4: support for a spiral-fold model of packaged DNA. *Proc. Natl. Acad. Sci. USA* **82**:7960–7964 (1985).
9. T. Blackstone, P. McGuirk, C. Laing, M. Vazquez, J. Roca and J. Arsuaga. The role of writhe in DNA condensation, Proceedings of International Workshop on Knot Theory for Scientific Objects. *OCAMI Studies* Osaka Municipal Universities Press, **1**:239–250 (2007).
10. F.P. Booy, W.W. Newcomb, B.L. Truss, J.C. Brown, T.S. Baker and A.C. Steven. Liquid-crystalline, phage-like packing of encapsidated DNA in herpes simplex virus. *Cell* **64**:1007–1015 (1991).
11. D. Buck and E. Flapan. Predicting knot or catenane type of site-specific recombination products. *J. Mol. Biol.* **374**:1186–1199 (2007).
12. C.R. Cantor and P.R. Schimel. *Biophysical Chemistry Part III: The Behavior of Biological Macromolecules*, (Freeman, San Francisco, 1980).
13. S. Casjens. Principles of virion structure, function and assembly, 1–3. In:*Structural Biology of Viruses.* Eds. W. Chiu, R. M. Burnett and R. Garcea. (Oxford University Press, 1997).
14. M. Cerritelli, N. Cheng, A. Rosenberg, C. McPherson, F. Booy and A. Steven. Encapsidated conformation of bacteriophage T7 DNA. *Cell* **91**:271–280 (1997).
15. D.K. Chattoraj and R.B. Inman. Location of DNA ends in P2, 186, P4 and lambda bacteriophage heads. *J. Mol. Biol.* **87**:11–22 (1974).
16. L.R. Comolli, A.J. Comolli, C.E. Spakowitz, P.J. Siegerist, P.J. Jardine, S. Grimes, D.L. Anderson, C. Bustamante and K.H. Downing. Three-dimensional architecture of the bacteriophage phi29 packaged genome and elucidation of its packaging process. *Virology* **371**:267–277 (2008).
17. I.K. Darcy, J. Chang, N. Druivenga, C. McKinney, R.K. Medikonduri, S. Mills, J. Navarra-Madsen, A. Ponnusamy, J. Sweet and T. Thompson. Coloring the Mu transpososome. *BMC Bioinformatics* **7**:435 (2006).
18. Y. Diao. The knotting of equilateral polygons in R^3. *J. Knot Theory and its Ramifications* **4**:189–196 (1995).
19. Y. Diao, A. Dobay, R.B. Kusner and A. Stasiak. The average crossing number of equilateral random polygons. *J. Phys. A* **36**:11561–11574 (2003).
20. R.W. Diebler, J.K. Mann, D.W. Sumners and L. Zechiedrich. Hin-mediated DNA knotting and recombining promote replicon dysfunction and mutation. *BMC Mol. Biol.* **8**:44 (2007).
21. T. Dokland, B.H. Lindqvist and S.D. Fuller. Image reconstruction from cryo-electron micrographs reveals the morphopoietic mechanism in the P2–P4 bacteriophage system. *EMBO J.* **3**:839–846 (1992).

22. W.C. Earnshaw and S.R. Casjens. DNA packaging by the double-stranded DNA bacteriophages. *Cell* **21**:319–331 (1980).
23. W.C. Earhshaw and S.C. Harrison. DNA arrangement in isometric phage heads. *Nature* **268**:598–602 (1977).
24. C. Ernst and D.W. Sumners. A calculus for rational tangles: applications to DNA recombination. *Math. Proc. Camb. Phil. Soc.* **108**:489–515 (1990).
25. A. Evilevitch, L. Lavelle, C.M. Knobler, E. Raspaud and W.M. Gelbart. Osmotic pressure inhibition of DNA ejection from phage. *Proc. Natl. Acad. Sci. USA* **100**: 9292–9295 (2003).
26. A. Flammini, A. Maritan and A. Stasiak. Simulations of action of DNA topoisomerases to investigate boundaries and shapes of spaces of knots. *Biophys. J.* **87**:2968–2975 (2004).
27. M.D. Frank-Kamenetskii, A.V. Lukashin and A.V. Vologodskii, AV. Statistical mechanics and topology of polymer chains. *Nature* **258**:398–402 (1975).
28. D.N. Fuller, J.P. Rickgauer, P.J. Jardine, S. Grimes, D.L. Anderson and D.E. Smith. Ionic effects on viral DNA packaging and portal motor function in bacteriophage phi 29. *Proc. Natl. Acad. Sci. USA* **104**:11245–11250 (2007).
29. S. Garoufalidis and P. Teichner. On knots with trivial Alexander polynomial. *J. Differential Geom.* **67**:167–193 (2004).
30. M. Gellert and D.R. Davies. Organization of DNA in bacteriophage P4. *J. Mol. Biol.* **8**:341–347 (1964).
31. I. Grainge, M. Bregu, M. Vazquez, V Sivanathan, S.C. Ip and D.J. Sherratt. Unlinking chromosomes catenated in vivo by site-specific recombination. *EMBO J.* **126**:4228–4238 (2007).
32. J. Hoste and M. Thistlethwaite. KNOTFIND, (1999), www.math.utk.edu/morwen/knotscape.html.
33. X. Hua, B. Baghavan, D. Nguyen, J. Arsuaga and M. Vazquez. Random state transitions of knots: a first step towards modeling unknotting by type II topoisomerases. *Topology and its Applications* **154**:1381–1397 (2007).
34. N. Hud. Double-stranded DNA organization in bacteriophage heads: an alternative toroid-based model. *Biophys. J.* **69**:1355–1362 (1995).
35. N.V. Hud and I.D. Vilfan. Toroidal condensates: unraveling the fine structure and the role of nucleation in determining size. *Annual Reviews* **34**:295–318 (2005).
36. B. Ibarra, J.R. Caston, O. Llorca, M. Valle, J.M. Valpuesta and J.L. Carrascosa. Topology of the components of the DNA packaging machinery in the phage Phi29 prohead. *J. Mol. Biol.* **298**:807–815 (2000).
37. M. Isaksen, B. Julien, R. Calendar and B.H. Lindquist. DNA topoisomerase protocols. In: *DNA Topology and Enzymes.* Eds. M.A. Bjornsti, N. Osheroff. (Humana, Totowa, NJ), **94**:69–74 (1999).
38. F. Jaeger, D. Welsh and D. Vertigan. On the computational complexity of the Jones and Tutte polynomials. *Proc. Cambridge Phil. Soc.* **108**:5–53 (1990).
39. W. Jiang, J. Chang, J. Jakana, P. Weigele, J. King and W. Chiu. Structure of epsilon15 bacteriophage reveals genome organization and DNA packaging/injection apparatus. *Nature* **439**:612–616 (2006).

40. J.E. Johnson and W. Chiu. Structures of virus and virus-like particles. *Curr. Op. in Str. Biol.* **10**:229–235 (2000).
41. J. Kindt, S. Tzlil, A. Ben-Shaul and W.M. Gelbart. DNA packaging and ejection forces in bacteriophage. *Proc. Natl. Acad. Sci. USA* **98**:13671–13674 (2001).
42. E. Kellenberger, E. Carlemalm, J. Sechaud, A. Ryter and G. Haller. Considerations on the condensation and the degree of compactness in non-eukaryotic DNA-containing plasmas, In: *Bacterial Chromatin: Proceedings of the Symposium "Selected Topics on Chromatin Structure and Function"*. Eds. C. Gualerzi and C.L. Pon. (Springer, Berlin, 1986), 11–25.
43. L.D. Kosturko, M. Hogan and N. Dattagupta. Structure of DNA within three isometric bacteriophages. *Cell* **16**:515–522 (1979).
44. K. Kimura, V.V. Rybenkov, N.J. Crisona, T. Hirano and N.R. Cozzarelli. 13S condensin actively reconfigures DNA by introducing global positive writhe: implications for chromosome condensation. *Cell* **98**:239–248 (1999).
45. K.V. Klenin, A.V. Vologodskii, V.V. Anshelevich, A.M. Dykhne and M.D. Frank-Kamenetskii. Effect of excluded volume on topological properties of circular DNA. *J. Biomol. Struct. Dyn.* **5**:1173–1185 (1988).
46. J.C. LaMarque, T.L. Le and S.C. Harvey. Packaging double-helical DNA into viral capsids. *Biopolymers* **73**:348–355 (2004).
47. J. Lepault, J. Dubochet, W. Baschong and E. Kellenberger. Organization of double-stranded DNA in bacteriophages: a study by cryo-electron microscopy of vitrified samples. *EMBO J.* **6**:1507–1512 (1987).
48. L.F. Liu, J.L. Davis and R. Calendar. Novel topologically knotted DNA from bacteriophage P4 capsids: studies with DNA topoisomerases. *Nucleic Acids Res.* **9**:3979–3989 (1981a).
49. L.F. Liu, L. Perkocha, R. Calendar and J.C. Wang. Knotted DNA from bacteriophage capsids. *Proc. Natl. Acad. Sci. USA* **78**:5498–5502 (1981b).
50. Z. Liu, E.L. Zechiedrich and H.S. Chan. Inferring global topology from local juxtaposition geometry: interlinking polymer rings and ramifications for topoisomerase action. *Biophys. J.* **90**:2344–2355 (2006).
51. C.R. Locker, S.D. Fuller and S.C. Harvey. DNA organization and thermodynamics during viral packing. *Biophys J.* **93**:2861–2869 (2007).
52. M.L. Mansfield. Knots in Hamilton cycles. *Macromolecules* **27**:5924–5926 (1994).
53. J. Menissier, G. de Murcia, G. Lebeurier and L. Hirth. Electron microscopic studies of the different topological forms of the cauliflower mosaic virus DNA: knotted encapsidated DNA and nuclear minichromosome. *EMBO J.* **2**: 1067–1071 (1983).
54. J.P.J. Michels and F.W. Wiegel. On the topology of a polymer ring. *Proc. R. Soc. London Ser. A* **403**:269–284 (1986).
55. C. Micheletti, D. Marenduzzo, E. Orlandini and D.W. Sumners. Knotting of random ring polymers in confined spaces. *J. Chem. Phys.* **124**:64903 (2006).
56. C. Micheletti, D. Marenduzzo, E. Orlandini and D.W. Sumners. Simulations of knotting in confined DNA rings. *Biophys J.* **95**:3591–3599 (2008).

57. K. Millett. Knotting of regular polygons in 3-space. In: *Random Knotting and Linking*. Eds. K. Millet and D. Sumners. (World Scientific, Singapore, 1994), 31–46.
58. K. Millett. Monte Carlo explorations of polygonal knot spaces. In: *Knots in Hellas '98* (World Scientific, Singapore, 2000), 306–334.
59. J.E. Mueller, S.M. Du and N.C. Seeman. The design and synthesis of a knot from single-stranded DNA. *J. Am. Chem. Soc.* **113**:6306–6308 (1991).
60. K. Murasugi. *Knot Theory and its Applications.* (Birkhauser, 2007).
61. L. Olavarrieta, M.L. Martinez-Robles, P. Hernandez, D.B. Krimer and J.B. Schvartzman. Knotting dynamics during DNA replication. *Mol. Microbiol.* **46**:699–707 (2002).
62. S. Pathania, M. Jayaram and R.M. Harshey. Path of DNA within the Mu transpososome. Transposase interactions bridging two Mu ends and the enhancer trap five DNA supercoils. *Cell* **109**:425–36 (2002).
63. A.S. Petrov, M.B. Boz and S.C. Harvey. The conformation of double-stranded DNA inside bacteriophages depends on capsid size and shape. *J. Struct. Biol.* **160**:241–248 (2007).
64. G.R. Pruss, J.C. Wang and R. Calendar. In vitro packaging of satellite phage P4 DNA. *Proc. Natl. Acad. Sci. USA* **71**:2367–2371 (1974).
65. E.J. van Rensburg, E. Orlandini, D.W. Sumners, M.C. Tesi and S.G. Whittington. The writhe of a self-avoiding polygon. *J. Phys A* **26**:L981–L986 (1994).
66. K. Richards, R. Williams and R. Calendar. Model of DNA packing within bacteriophage heads. *J. Mol. Biol.* **78**:255–259 (1973).
67. S. Rishovd, A. Holzenburg, B.V. Johansen and B.H. Lindquist. Bacteriophage P2 and P4 morphogenesis: structure and function of the connector. *Virology* **245**:11–17 (1998).
68. V.V. Rybenkov, N.R. Cozzarelli and A.V. Vologodskii. Probability of DNA knotting and the effective diameter of the DNA double helix. *Proc. Natl. Acad. Sci. USA* **90**:5307–5311 (1993).
69. V.V. Rybenkov, C. Ullsperger, A.V. Vologodskii and N.R. Cozarelli. Simplification of DNA topology below equilibrium values by type II topoisomerases. *Science* **27**:690–693 (1997).
70. J. Roca. Varying levels of positive and negative supercoiling differently a?ect the efficiency with which topoisomerase II catenates and decatenates DNA. *J. Mol. Biol.* **305**:441–450 (2001).
71. R. Scharein. *Interactive Topological Drawing*. PhD Thesis, Department of Computer Science, University of British Columbia, 1998.
72. M. Schmutz, D. Durand, A. Debin, Y. Palavadeau, E.R. Etienne and A.R. Thierry. DNA packing in stable lipid complexes designed for gene transfer imitates DNA compaction in bacteriophage. *Proc. Natl. Acad. Sci. USA* **96**:12293–12298 (1999).
73. N.C. Seeman. Biochemistry and structural DNA nanotechnology: an evolving symbiotic relationship. *Biochemistry* **42**:7259–7269 (2003).
74. P. Serwer. Arrangement of double-stranded DNA packaged in bacteriophage capsids: an alternative model. *J. Mol. Biol.* **190**:509–512 (1986).

75. P. Serwer, S.J. Hayes and R.H. Watson. Conformation of DNA packaged in bacteriophage T7. Analysis by use of ultraviolet light-induced DNA-capsid cross-linking. *J. Mol. Biol.* **223**:999–1011 (1992).
76. S.Y. Shaw and J.C. Wang. Knotting of a DNA chain during ring closure. *Science* **260**:533–536 (1993).
77. K. Shishido, N. Komiyama and S. Ikawa. Increased production of a knotted form of plasmid pBR322 DNA in *Escherichia coli* DNA topoisomerase mutants. *J. Mol. Biol.* **195**:215–218 (1987).
78. J.L. Sikorav, J. Pelta and F. Livolant. A liquid crystalline phase in spermidine-condensed DNA. *Biophys. J.* **67**:1387–1392 (1994).
79. D.E. Smith, S.J. Tans, S.B. Smith, S. Grimes, D.L. Anderson and C. Bustamante. The bacteriophage phi29 portal motor can package DNA against a large internal force. *Nature* **413**:748–752 (2001).
80. T.E. Strzelecka, M.W. Davidson and R.L. Rill. Multiple liquid crystal phases of DNA at high concentrations. *Nature* **331**:457–60 (1988).
81. M.C. Tesi, E.J. Janse van Rensburg, E. Orlandini, D.W. Sumners and S.G. Whittington. Knotting and supercoiling in circular DNA: a model incorporating the effect of added salt. *Phys. Rev. E* **49**:868–872 (1994).
82. S. Trigueros, J. Arsuaga, M. Vazquez, D.W. Sumners and J. Roca. Novel display of knotted DNA molecules by two dimensional gel electrophoresis. *Nucleic Acids Research* **29**:E67 (2001).
83. S. Trigueros and J. Roca. Production of highly knotted DNA by means of cosmid circularization inside phage capsids. *BMC Biotechnol.* **7**:94 (2007).
84. S. Tzlil, J.K. Kindt, W.M. Gelbart and A. Ben-Shaul. Forces and pressures in DNA packaging and release from viral capsids. *Biophys. J.* **84**:1616–1627 (2003).
85. R. Varela, K. Hinson, J. Arsuaga and Y. Diao. A fast ergodic algorithm for generating ensembles of random polygons. *J. Phys A: Math Gen* **42**:095204 (2009).
86. M. Vazquez and D.W. Sumners. Tangle analysis of gin site-specific recombination. *Math. Proc. Camb. Phil. Soc.* **136**:565–582 (2004).
87. A.A. Vetcher, A.Y. Lushnikov, J. Navarra-Madsen, R.G. Scharein, Y.L. Lyubchenko, I.K. Darcy and S.D. Levene. DNA topology and geometry in Flp and Cre recombination. *J. Mol. Biol.* **357**:1089–1104 (2006).
88. M. Vazquez, S.D. Colloms and D.W. Sumners. Tangle analysis of Xer recombination reveals only three solutions, all consistent with a single three-dimensional topological pathway. *J. Mol. Biol.* **346**:493–504 (2005).
89. A.V. Vologodskii, N.J. Crisona, B. Laurie, P. Pieranski, V. Katritch, J. Dubochet and A. Stasiak. Sedimentation and electrophoretic migration of DNA knots and catenanes. *J. Mol. Biol.* **278**:1–3 (1998).
90. A.V. Vologodskii, W. Zhang, V.V. Rybenkov, A.A. Podtelezhnikov, D. Subramanian, J.D. Griffith and N.R. Cozzarelli. Mechanism of topology simplification by type II DNA topoisomerases. *Proc. Natl. Acad. Sci. USA* **98**: 3045–3049 (2001).
91. S. Wang, J.R. Chang and T. Dokland. Assembly of bacteriophage P2 and P4 procapsids with internal scaffolding protein. *Virology* **348**:133–140 (2006).

92. J.C. Wang, K.V. Martin and R. Calendar. On the sequence similarity of the cohesive ends of coliphage P4, P2, and 186 deoxyribonucleic acid. *Biochemistry* **12**:2119–2123 (1973).
93. S.A. Wasserman and N.R. Cozzarelli. Biochemical topology: applications to DNA recombination and replication. *Science* **232**:951–960 (1986).
94. S.A. Wasserman and N.R. Cozzarelli. Supercoiled DNA-directed knotting by T4 topoisomerase. *J. Biol. Chem.* **266**:20567–20573 (1991).
95. C. Weber, A. Stasiak, P. de los Rios and G. Dietler. Numerical simulation of gel electrophoresis of DNA knots in weak and strong electric fields. *Biophys. J.* **90**:3100–3105 (2006).
96. J.S. Wolfson, G.L. McHugh, D.C. Hooper and M.N. Schwartz. Knotting of DNA molecules isolated from deletion mutants of intact bacteriophage P4. *Nucleic Acids Res.* **13**:6695–6702 (1985).
97. E.L. Zichiedrich and N.J. Crisona. Coating DNA with RecA protein to distinguish DNA path by electron microscopy. In: *DNA Topoisomerase Protocols, DNA Topology and Enzymes*. Eds. M.A. Bjornsti, N. Osheroff. (Humana, Totowa, NJ), **94**:99 (1999).